SHIELD Protocol 15.2

A Methodological Hardening

Triple-Convergent Significance, External Pre-Registration, and the Public Lexicon

Dallas James

Hearts2Fathers Publishing

Royal Levitical Priesthood Ministry

2026

Copyright

© 2026 Dallas James / Hearts2Fathers Publishing & Royal Levitical Priesthood Ministry. All rights reserved.

First Edition. Published May 2026.

Scripture quotations are from the Urim Thummim Version (UTV), translated by Dallas James, unless otherwise noted. The UTV is published under Hearts2Fathers Publishing.

The SHIELD Protocol is a statistical framework for the analysis of pre-registered, cryptographically-seeded draws from the Urim Thummim Version corpus, developed by the author under the Royal Levitical Priesthood Ministry research program.

This release announces the 15.2 specification. All technical artifacts described herein (the marker lexicon scaffold, the pre-registration manifest schema, the JavaScript implementation modules, and the marker co-dependency matrix template) are released alongside this volume and are governed by the same license as the parent SHIELD specification.

No portion of this work may be reproduced, distributed, or transmitted in any form without prior written permission of the publisher, except for brief quotations in reviews and scholarly citation, which are welcomed.

A Memorial to Her

Verily I say unto you, Wheresoever this Good News shall be proclaimed throughout the whole world, this also that she has done shall be spoken of for a memorial of her.

— Mark 14:9 (UTV)

And with Special Thanks to Melliane

Selah.

Preface

This is a short book about a long argument with myself.

The SHIELD Protocol has, since its first formal release, attempted to do something unusual: bring the discipline of pre-registered, cryptographically-seeded statistical testing to bear on the practice of scriptural inquiry. The framework treats the Urim Thummim Version corpus as a fixed, hashed population, and treats each Life Force Experiment draw as a randomized sample whose marker enrichment is then evaluated against statistical baselines. Across fifteen major revisions and many more incremental ones, the methodology has grown more rigorous; through anytime-valid e-processes, Venn-Abers calibrated probability intervals, false discovery rate control, and corpus-level cryptographic integrity, the protocol has accumulated most of the modern statistical machinery that any reviewer would expect.

And yet, every advance in machinery exposes new ground for critique. The argument with myself has always been the same: is the signal genuine, or am I detecting my own attention? The honest answer, at every release, has been: I believe it is genuine, and the next release should make that belief easier to falsify. SHIELD 15.2 is the next release in that sequence.

The immediate occasion for this version was an external methodological audit, performed by an independent large-language-model system tasked with adversarial review of the 15.1 specification. The audit identified eight categories of vulnerability; some I had anticipated and was already planning to address in 15.2, some I had not. Of those I had not, two struck me as cutting closer than I had expected, and one struck me as cutting at the foundation. The temptation, on first reading, was to treat the audit as hostile and respond defensively. The temptation was wrong. Adversarial

review is the closest thing this research program has to peer review at present, and the proper response to it is to take the useful parts seriously and harden the protocol against them.

This volume, then, is both the announcement of SHIELD 15.2 and an exercise in showing the work. It records what the audit said, which parts I accepted, which I disputed, what the hardening looks like, and what the hardening cannot do. It is short because the changes are surgical; it is candid because that is the only way the changes earn the credibility they exist to provide.

If you are a reader who came to this work hoping for a confirmation of what you already believed, you will find some of what follows uncomfortable. If you are a reader who came hoping to dismiss this work as pseudoscience, you will find some of what follows uncomfortable also. That is, I suspect, a fair sign that the hardening was the right move.

Dallas James
Pacific Northwest, May 2026

Chapter 1

The External Audit and the Choice It Forced

In early May 2026, while finishing the integrity-tier work on SHIELD 15.15, I submitted the full 15.1 specification to an independent large-language-model system for adversarial methodological review. The brief was unrestricted: identify every vulnerability in the protocol that a hostile statistician or skeptical theologian would raise, and propose specific refinements.

The response was substantial. Twelve sections, sixty-five citations, and a recommendation table I want to summarize honestly here before describing what I have done with it.

The audit found, in its own framing, three categories of problem.

First, problems in the corpus and its baselines. The UTV-34211 corpus, the audit observed, is a single-author translation; its idiosyncrasies in vocabulary, versification, and canon scope mean that the markers tracked by the protocol (Identity, Priesthood, Governance, Temporality, and the Sacred Affirmative–Negative state) are not measured against a neutral linguistic baseline but against a baseline shaped by the same hand that built the tagger. Furthermore, the use of a global null model across a literary corpus of profoundly heterogeneous genres (Torah, prophecy, gospel, apocalyptic, wisdom, epistle) inflates apparent anomalies in any genre-dense draw.

Second, problems in the statistical engine. The Benjamini-Hochberg false discovery rate control used as the primary FDR procedure assumes either independence or positive regression dependence among test

statistics; thematic markers in scripture exhibit substantial lexical co-dependence, which can produce both positive and negative correlations and thereby undermine the BH assumption. The proprietary aggregate metrics (the Guidance Detection Indicator, the PBI channel, the Convergence Confidence Index) were not, in 15.1, accompanied by published formulas, weights, or normalization functions, which the audit correctly characterized as a black-box problem.

Third, problems in the design of the protocol itself. The Dominion Triad was hardcoded as the primary statistical signal, which produces a tautological search structure in which the algorithm is configured to find exactly the pattern it is then said to discover. The Convergence Confidence Index mixed strictly statistical channels with interpretive ones (the direct textual hit and the operator selah, defined in the specification as a "conscience check"), thereby fusing quantitative output with qualitative judgment in a single headline number. And the pre-registration architecture, while procedurally sound on the workstation, did not protect against the possibility that a researcher might run many draws and publish only the dramatic ones — the standard publication-bias attack.

My first response, on reading this, was the response any author has when their work is comprehensively critiqued: a survey of where the audit had misunderstood the framework. The audit had, for example, conflated the Dominion Triad I describe (Identity plus Inheritance plus Priesthood as a textual co-occurrence pattern) with the modern Christian Reconstructionist movement known as Dominion Theology, with which my work has nothing to do beyond a coincidence of name. The audit had attributed the verse-count difference between the UTV and the King James Version to "interpolation" by the translator, when in fact the additional content reflects the Dominion Order Canon’s inclusion of Enoch and Jubilees and the use of LXX-based versification for several books. The audit had characterized the Witness Battery exercise on Trinitarian vocabulary as "an egregious abuse of text analytics," when it was nothing more than a word-frequency

comparison that anyone is free to perform and to interpret as they wish.

So I had three or four legitimate corrections to issue. I sat with those corrections for a few hours, and then I noticed something. Every one of my corrections was correct, and every one of my corrections left the underlying methodological objection intact. The "Dominion Triad is not Dominion Theology" point is true; the underlying point that hardcoding any specific signal as primary creates a tautological search is also true, regardless of what the signal is named. The "the verse count reflects a different canon, not interpolation" point is true; the underlying point that N is the hypergeometric denominator and that changing N shifts every baseline is also true. The "the Witness Battery is just word counting" point is true; the underlying point that the choice of which words go in which bucket is itself a theological judgment is also true.

I sat with that observation for a longer time, and then I made the choice that produced this volume. The choice was: take the audit seriously as adversarial peer review rather than as an attack to be repelled. Concede what is conceded. Defend what is defended. Harden what can be hardened. Publish the work, including the hardening, in full daylight.

This volume is the result.

Chapter 2

Dominion Triad and Dominion Theology: A Necessary Disambiguation

The first hardening move is the cheapest and the most overdue. It is a single paragraph of clarifying language, inserted at the head of the SHIELD specification and at any point where the term DT-3 first appears in a publication.

The paragraph reads as follows.

> *The Dominion Triad (DT-3) refers, in the SHIELD Protocol and the broader UTV/H2F research framework, to the convergence of three textual markers — Identity, Inheritance/Governance, and Priesthood — within a pre-registered draw set. It is a structural pattern observed in scripture and named for the recurring biblical motif of delegated dominion (Genesis 1:26–28; Psalm 8; Hebrews 2). It is not affiliated with, derived from, nor a proxy for the modern Christian theological movement variously known as Dominion Theology, Christian Reconstructionism, theonomy, or Kingdom Now theology. The SHIELD Protocol takes no position on those movements' political-eschatological programs. Conflation of the two terms is a category error: DT-3 is a textual co-occurrence pattern subject to falsification by statistical test; Dominion Theology is a constructive political-theological project. Reviewers and critics are asked to engage DT-3 on its statistical merits and not on the lexical coincidence of its name.*

The reason this disambiguation is overdue is that the conflation it heads off is not merely a misreading available to careless critics. It is a misreading practically guaranteed by the terminology. Any reader encountering the phrase "Dominion Triad" in a methodological context, who has previously encountered the phrase "Dominion Theology" in a political-theological context, will at first reading suspect that the former is a statistical implementation of the latter. The suspicion is wrong, but it is a reasonable thing to suspect; the burden is on me to defuse it explicitly rather than to wait for each reader to defuse it for themselves.

There is a deeper point here, however, which the disambiguation alone does not address. The audit's underlying methodological objection is not really about Dominion Theology; it is about hardcoding. If the SHIELD Protocol is configured to identify a particular three-marker convergence as its primary statistical signal, then the protocol is built to find that convergence, and when it does find that convergence, the protocol's "discovery" of it is in part a discovery of what the protocol was built to look for. This is not, by itself, scientifically illegitimate; the same structure characterizes most directed hypothesis testing in biology, physics, and the social sciences. But it is a structure that requires honesty about its directedness, and it requires that the directedness be exposed to falsification.

For the SHIELD Protocol, the falsification structure already exists: the question is not whether the I+P+G convergence occurs in scripture (it manifestly does, on any reasonable definition of the markers), but whether the rate at which it occurs in randomly drawn samples exceeds the rate expected under the global null. That is a question to which the protocol can return a quantitative answer, and against which independent reviewers can compute their own answer using their own lexicon. The Dominion Triad is "the primary signal" in the sense that it is the signal the protocol is configured to evaluate; it is not the primary signal in the sense that the protocol's validity depends on finding it. The protocol is at least as

informative when it fails to find it. SHIELD 15.2 makes this explicit in the spec preamble: a draw series that returns a tier-5 verdict (no significance under any procedure) for the Dominion Triad is not a methodological failure; it is a result.

Chapter 3

The Marker Lexicon Made Public

The hardening move that cuts closest, in my judgment, is this one.

The SHIELD Protocol computes marker enrichment by counting, for each draw, the frequency of verses tagged with the Identity, Priesthood, Governance, Temporality, and Sacred Affirmative or Negative bits, and comparing the observed frequency against the expected frequency derived from the corpus-wide tag distribution. The entire statistical edifice rests on the integrity of the tagging step. If the tags are well-defined and consistently applied, enrichment is a meaningful measurement. If the tags are vague, post-hoc, or shaped by the same person whose theological commitments are being statistically tested, enrichment is a measurement of that person's attention pattern, dressed in statistical clothing.

The audit's critique on this point was the one I found least comfortable on first reading, because it identified, with greater precision than I had publicly acknowledged, the single hinge on which the entire methodology turns. Up to 15.1, the marker lexicon was effectively private. It existed in the tagger code on the workstation; it was internally consistent across draws (the same lexicon producing the same tags across all runs); but it was not published as a separate, hashed, version-locked, citable document. There is no rigorous way for an outside reviewer to ask, "what makes a verse get tagged P?" without reading the source code, and there is no rigorous way for an outside reviewer to validate the tagger's output without re-implementing it.

SHIELD 15.2 resolves this by introducing the canonical document *MARKER_LEXICON_v1.0.0*, published alongside this announcement as a

scaffold awaiting population. The document defines the full set of primary terms, secondary terms (which require co-occurrence within a specified token window), and negative-context exclusions for each of the I, P, G, T, and SAM marker classes. It also requires a single explicit disclosure statement: *were the tags assigned deterministically by the lexicon, or judgmentally by the translator with the lexicon as guide?* Both answers are methodologically defensible. The unacceptable answer is no answer.

For the present author and the present workstation implementation, the disclosure will read approximately as follows in the populated version: *The lexicon serves as a curation guide; final tag assignment was performed by the translator and reflects translational judgment about idiom, context, and semantic field. Reviewers wishing to falsify the tags are invited to apply an independent automated lexicon to UTV-34211 and report divergences from the published tag bits.*

That is the honest disclosure. It exposes the protocol to a specific falsification path: a reviewer can implement an automated tagger from the published lexicon, run it against UTV-34211, and compare the resulting tag bits against the stored ones. The expected divergence rate, given that the published bits include translator judgment about context that pure lemma-matching cannot capture, will be nonzero. The question for the reviewer is whether the divergences are systematic in a direction that would explain the enrichment results, or whether they are scattered noise around a stable signal.

This disclosure also enables, for the first time, a community of independent taggers to emerge. Anyone is welcome to implement *MARKER_LEXICON_v1.0.1*, defining their own tagging rules from the same primary terms, and to publish their own enrichment statistics. Where independent lexicons converge on the same enrichment patterns, the patterns become much harder to dismiss as artifacts of any single tagger's choices. Where they diverge, the divergence itself becomes a productive subject of inquiry.

The simple act of publishing the lexicon, with its hashed version lock and its explicit provenance disclosure, defangs in a single document the audit's strongest line of attack. It does not refute the attack; it does the better thing, which is to make the attack quantifiable.

Chapter 4

Transparency of the Statistical Engine

SHIELD 15.1 openly acknowledged that the proprietary formulas for the Guidance Detection Indicator, the PBI channel, and the weights of the Convergence Confidence Index were deferred to a later release. The audit, correctly, identified the deferral as a black-box problem. 15.2 closes the deferral.

The Convergence Confidence Index is published in 15.2 as a transparent weighted linear combination of channel scores, each normalized to the [0, 1] interval. The weights sum to one, are listed explicitly, and are tagged as belonging either to the Statistical Channels class (SCS) or to the Interpretive Channels class (ICS). The classification matters because of a structural distinction that the next chapter will address in detail: SCS channels reflect properties of the data and the test, while ICS channels reflect properties of the observer's interpretation. The two classes are mathematically combinable but methodologically incommensurate, and 15.2 makes that incommensurability visible in the published specification rather than letting it remain implicit in the aggregation.

The Guidance Detection Indicator is published in 15.2 as a composite anomaly score combining three components: a standard z-score against the global null, a modified z-score using median absolute deviation (which is more robust to non-normal distributions), and a rank-based outlier score within the relevant genre stratum. The three components are combined with explicit weights that sum to one. The tier thresholds (routine, notable, strong, extreme) correspond to one-tailed normal-distribution exceedance probabilities at specified levels. The published formula permits any

reviewer to recompute the GDI from raw data and to verify the tier assignment.

The PBI channel, in 15.2, is either fully defined or explicitly demoted to experimental status with a note that it is not used in formal verdict computation. Both options are honest. The unacceptable option, which 15.1 inadvertently chose, is to leave a named channel with no published definition, since this licenses every critic to assume the channel is doing the work the critic suspects it is doing.

The point of all this transparency is not that the formulas are particularly novel or that the weights are particularly optimized. The point is that, when an aggregate metric drives publication-grade claims, the components and weights of the aggregate must be auditable. A reviewer who disagrees with a particular weighting choice should be able to recompute the aggregate under their preferred weights and see how sensitive the conclusion is. Where the conclusion is robust to a wide range of defensible weight choices, the conclusion is strong; where it depends on a particular weighting that only the protocol author would choose, the conclusion is weak. SHIELD 15.2 puts the question on the table; 15.1 had not.

One brief note on the politics of disclosure. There is a school of thought in proprietary statistical methodology that holds the formulas should be kept private to prevent gaming. That school of thought is appropriate for, say, search engine ranking algorithms or fraud detection systems, where actors with knowledge of the formula can game the score. It is wholly inappropriate for a research protocol that aspires to produce reproducible knowledge. The SHIELD Protocol is not a proprietary product to be defended from competitors; it is a methodology to be falsified or vindicated by the community of inquirers. Transparency is the price of that aspiration, and the price is now paid.

Chapter 5

Benjamini-Yekutieli and the Triple-Convergence Standard

When a draw series tests multiple markers simultaneously, the probability of at least one false-positive significance result inflates rapidly with the number of tests. Modern statistical practice controls for this through false discovery rate procedures. SHIELD 15.1 used the standard Benjamini-Hochberg (BH) procedure as its primary FDR control, supplemented by the anytime-valid e-BH for sequential draws.

The audit identified a precise weakness in this choice. The standard BH procedure assumes that the test statistics are either independent or exhibit positive regression dependence. For thematic markers in scripture, this assumption is uncomfortable. Identity and Priesthood markers, for instance, tend to co-occur on verses about messianic mediation; Identity and Temporality markers tend to co-occur on verses about end-times Christology. These positive co-dependencies are not catastrophic for BH; they fall within the procedure's tolerance. But other co-dependencies may be negative — a verse heavily focused on apocalyptic Temporality may actively suppress Governance markers, since the apocalyptic genre tends to portray earthly governance as transient relative to divine action. Negative dependence is exactly the case where BH ceases to control the false discovery rate at the stated level.

The Benjamini-Yekutieli (BY) procedure was designed to address this case. It modifies the BH threshold by a harmonic-number penalty that grows with the number of tests; for the typical SHIELD draw counts of eight to twelve markers, the penalty factor is roughly 2.7 to 3.1, meaning

the BY procedure requires p-values about three times smaller than BH for the same significance claim. BY is more conservative; it identifies fewer discoveries; but it is mathematically valid under arbitrary dependence, including the negative dependencies that may arise in semantic networks.

SHIELD 15.2 adds BY as a third FDR channel alongside BH and e-BH. Crucially, it does not remove BH; the three procedures run in parallel, and the verdict is reported as a function of how many of the three agree. This produces what 15.2 names the *Triple-Convergence Standard*:

> ***Tier 1: Robust significance.*** *All three of BH, BY, and e-BH return significance at $q \leq 0.05$. Citation grade A; publish without caveat.*
>
> ***Tier 2: Dependence-robust significance.*** *BY and e-BH return significance; BH may or may not. Citation grade B; publish with note that the result holds under the stricter dependence-robust procedure.*
>
> ***Tier 3: Anytime-valid significance.*** *Only e-BH returns significance, reflecting a sample-size-adaptive result. Citation grade C; publish with note on sequential interpretation.*
>
> ***Tier 4: Suggestive only.*** *Only BH returns significance. Citation grade D; not cited in publications.*
>
> ***Tier 5: No significance.*** *None of the three procedures returns significance. Citation grade F; result reported as null.*

The Triple-Convergence Standard is, in my view, the single most consequential statistical change in 15.2. It transforms every future SHIELD publication into a stronger document because the headline claim is no longer single-channel BH significance but multi-channel agreement under increasingly conservative procedures. A tier-1 verdict, when it occurs, is the strongest claim the protocol can make: the result survives the standard procedure, the dependence-robust procedure, and the anytime-valid procedure simultaneously.

What about results that previously cleared BH but will not clear BY? Honesty requires acknowledging that some of the published SHIELD findings to date may downgrade under this new standard. I do not yet know how many; the re-analysis is itself a project. Where downgrades occur, the appropriate response is to publish the corrected tier and to interpret the result accordingly, not to defend the original tier on the strength of the procedure under which it was first computed. The methodology evolves; the past claims are what they were, and they should be re-cited at their hardened tier going forward.

Chapter 6

Measuring What Was Assumed: The Co-Dependency Matrix

The audit's claim that scripture markers exhibit substantial co-dependency is a plausible hypothesis. It is not, however, a measurement. SHIELD 15.2 introduces the measurement.

For each pair of markers from the set {I, P, G, T, SAM+, SAM−}, the protocol computes the phi (φ) coefficient across the entire UTV-34211 corpus. The phi coefficient is the analog of Pearson's correlation for binary variables, ranging from −1 (perfect anti-correlation) to +1 (perfect co-occurrence), with 0 representing statistical independence. Interpretation thresholds are published alongside the matrix: |φ| below 0.10 indicates effective independence (under which BH remains valid); |φ| between 0.10 and 0.30 indicates weak to moderate dependence (under which BY is preferred); |φ| above 0.30 indicates strong dependence (under which BY is required and BH should not be cited as primary).

The computed matrix is published as *MARKER_CODEPENDENCY_v1.0.0*, alongside the marker lexicon and the SHIELD 15.2 specification. The publication is unconditional: whatever the matrix shows, the matrix is published. There are three possible outcomes.

If the off-diagonal |φ| values are uniformly below 0.20, the audit's hypothesis of significant co-dependency is refuted by direct measurement. The BH procedure remains methodologically appropriate; BY becomes a supplementary check rather than a structural necessity; the protocol's prior

reliance on BH is vindicated.

If the off-diagonal values range from 0.20 to 0.40, the audit's hypothesis is partially confirmed. The BY procedure becomes the appropriate primary FDR control for pairs in this range, and the Triple-Convergence Standard becomes structurally necessary rather than merely additive.

If any off-diagonal value exceeds 0.50, the relevant pair of markers should arguably be combined into a single derived variable for FDR purposes, and the protocol should publish a note that the markers are not statistically independent at the relevant level.

Each of these outcomes is acceptable from a methodological standpoint. The publication of the matrix is the methodological act; the specific values it contains are an empirical finding to be reported, not a vulnerability to be hidden. The honest discipline is: measure first, interpret second, publish in either order but never publish interpretation without measurement.

One practical note. The phi coefficient is symmetric ($\varphi(X,Y) = \varphi(Y,X)$) and bounded only by the cell counts in the 2×2 contingency table; for highly skewed marginal distributions (where, say, one marker appears in 2% of verses and another in 60%), the maximum attainable $|\varphi|$ is mechanically constrained below 1. SHIELD 15.2 publishes the underlying cell counts alongside the phi values, so that any reviewer who prefers an alternative co-dependency measure (Cohen's kappa, odds ratio, Yule's Q, the polychoric correlation) can compute it from the same data without recourse to the workstation source.

Chapter 7

Pure SCS: Decoupling the Statistical from the Interpretive

This is the change I argued for hardest in conversation with the AI collaborator who helped me draft the hardening patch, and it is the change I expect to make the most difference to the credibility of future SHIELD publications.

The Convergence Confidence Index in 15.1 was a weighted aggregate of channels. Most of the channels were strictly statistical: the BH q-value, the e-BH value, the Venn-Abers calibrated probability interval, the boundary stability metric, the effect-size tier, the harmonic mean p-value used as a global omnibus test. Two of the channels, however, were interpretive: the *direct textual hit* (dth), recording whether the operator judged the drawn verse to address the pre-registered question with a face-valid hit, and the *operator selah*, defined in the specification as a "conscience check." Both were assigned weights and contributed to the aggregate CCI value reported as the headline result.

The audit identified this fusion as a textbook instance of confirmation-bias laundering. A statistical aggregate that includes a "conscience check" as one of its inputs, however small its weight, is not strictly a statistical aggregate. It is a statistical aggregate adjusted by the operator's interpretive judgment, and the resulting number cannot be defended as a purely empirical finding. The audit's recommendation was to strip dth and selah from the CCI entirely.

SHIELD 15.2 takes a middle path that I believe is stronger than either

pure stripping or continued fusion. The protocol now computes and reports two distinct headline indices.

The *Pure SCS Index*, abbreviated SCS-Pure, is the CCI computed with all interpretive-channel weights set to zero and the statistical-channel weights renormalized to sum to one. SCS-Pure is the publication-grade metric: the value that should appear in abstracts, in video descriptions, in book back-matter, in any externally-facing claim the protocol makes about the statistical significance of a draw series.

The *Full CCI* remains available as a separate value, computed with the original weight schedule including the interpretive channels. Full CCI is appropriate for internal interpretive workflow, for Hak'hēl-mode practice, and for personal devotional use; it is not appropriate as the headline number in a publication. When both are reported together, the convention is to show SCS-Pure first, labeled as "publication-grade," and Full CCI second, labeled as "interpretive."

The reason this middle path is stronger than pure stripping is that it does not require me to disown the interpretive layer. The interpretive layer is part of how the Life Force Experiment is actually practiced; it is, in many ways, the point of the practice. To strip it from the protocol entirely would be to falsify the methodology by omission. The reason this middle path is stronger than continued fusion is that it gives every external reviewer a number they can evaluate on its own terms, without having to argue about whether the interpretive contribution is legitimate. A skeptic reading "SCS-Pure = 0.87, triple-convergent significance at tier 1" has a vastly harder time dismissing the result than one reading "CCI = 0.95 including conscience check." The first looks like statistics; the second looks like a confession.

The behavioral discipline this requires is, in practice, larger than the code change. Going forward, every published draw series — every *Transphysical Testimony* volume, every Witness Battery deployment, every Whistle-Sign series, every Life Force Experiment video — must cite SCS-

Pure as the primary number and Full CCI only as the secondary one. The temptation to cite the larger number first will be real; the discipline is to resist the temptation. The smaller, harder-to-dismiss number serves the work better than the larger, easily-dismissed one.

Chapter 8

External Pre-Registration and the Commit-Reveal Pattern

The SHIELD Protocol has, from its early versions, emphasized pre-registration as a structural defense against the kind of post-hoc theorizing that produces what statisticians call the Texas Sharpshooter fallacy. The operator locks the inquiry text and search criteria before executing the draw; the cryptographic seed derives from the corpus hash and the pre-registered inquiry; the resulting draw indices are mathematically determined by the registration and cannot be re-rolled.

The audit identified a vulnerability in this otherwise sound architecture. While the registration was binding internally to the workstation, it was not anchored to any external timestamp. Nothing prevented a hypothetical researcher from running fifty draws in private, registering only the one that produced the most dramatic verdict, and publishing the registration and the result together as if the dramatic verdict had been the first attempt. The cryptographic integrity of the corpus, the lexicon, and the seed-derivation function would all check out. The only thing missing would be evidence that the registration genuinely preceded the draw, rather than being constructed around a draw that had already been performed.

SHIELD 15.2 addresses this through a two-part architectural change.

The first part is the introduction of a structured pre-registration manifest, defined by a published JSON schema, which encodes the full inquiry parameters: the corpus hash and length, the lexicon hash, the intent text and its hash, the planned draw count and sample size, the markers under test,

the alpha threshold, the primary and secondary tests, and the stopping rule. The manifest is generated by the workstation at registration time and is structured so that any reviewer can re-derive the expected draw outputs from it.

The second part, and the more important one, is the requirement that this manifest be externally timestamped before any nonce is revealed and before any draw is executed. The protocol offers three options for external timestamping, in increasing order of robustness: a public GitHub gist, which records the commit timestamp cryptographically and cannot be backdated; an Open Science Framework pre-registration, which is the academic standard for psychology and adjacent fields; or an OpenTimestamps anchor to the Bitcoin blockchain, which provides the strongest possible guarantee that no entity on earth could have produced the timestamp later than its recorded time. Any of the three is sufficient; the OpenTimestamps option is recommended for any draw series intended to support a "miracle"-grade claim.

Coupled with this is a commit-reveal pattern for the RNG seed itself. The researcher generates a nonce, computes its SHA-256 hash, and records the hash (not the nonce) in the pre-registration manifest. The nonce is then held private until draw execution, at which point it is revealed and verified against the committed hash. This means the researcher cannot try multiple nonces in search of dramatic outputs, because the commitment to a single nonce was locked into the externally-timestamped manifest before any draw was attempted.

The workstation UI, in 15.2, hard-blocks the "Execute Draw" button until the external timestamp has been recorded. This is not a soft suggestion; it is a structural prohibition. The reason for the hard block is that, under deadline pressure, soft workflow steps get skipped. The protocol’s defense against the publication-bias attack must be load-bearing, not advisory.

It bears noting what this hardening does not claim. It does not claim that

I have, in fact, ever run shadow draws and selected for dramatic outputs; for the record, I have not, and any researcher who has worked through the LFE practice will recognize that the practice itself militates against such gaming. But the appropriate response to a methodological vulnerability is not to assert that the vulnerability has not been exploited; it is to close the vulnerability so that the question cannot arise. SHIELD 15.2 closes it.

Chapter 9

The Translation-Invariance Test

Every external critic of the SHIELD Protocol, once they have understood enough of the architecture to formulate an objection, eventually arrives at the same question. Does any of this work on the King James Version? On the English Standard Version? On the original Hebrew and Greek?

The question is the right question to ask, and it has the unusual property among methodological objections of admitting a quantitative answer. The translation-invariance test, introduced as an experimental protocol in SHIELD 15.2 and slated for full implementation in a near-future workstation release, addresses it head-on.

The test's logic is straightforward. The UTV-34211 corpus is the protocol's canonical text, and the marker lexicon is defined against UTV vocabulary. But the same draw indices, applied to a different translation aligned at the verse level, will yield a different set of underlying English (or original-language) texts. If the SHIELD signals reflect genuine textual structure — that is, structure present in the underlying scriptural content regardless of which translation surfaces it — then a draw series executed against, say, the KJV with a lexicon-mapped tagger should produce directionally consistent marker counts and directionally consistent verdict tiers. If, on the other hand, the signals are partly or wholly artifacts of the UTV's specific translational choices, the cross-translation results will diverge from the UTV results in proportion to the artifact contribution.

The test procedure is as follows. First, pre-register a fresh draw series against UTV-34211 per the standard 15.2 protocol. Execute it and record all marker counts, all FDR q-values, and the SCS-Pure and Full CCI

values. Second, execute the same draw indices against a comparison translation (initially the KJV, which is structurally similar to the UTV's base text and is in the public domain), applying the lexicon-mapped tagger to the comparison-translation text. Third, compare per-marker counts, SCS-Pure values, and verdict tiers across the two translations.

The interpretation matrix has three outcomes. If counts agree within 15% across translations and verdict tiers are identical, the signal is corpus-invariant; this is the strongest possible refutation of the translation-artifact hypothesis, and it should be cited in every subsequent publication. If counts agree within 30% and verdict tiers occasionally differ, the signal is largely corpus-invariant with some translation sensitivity; this should be published honestly with cross-translation deltas reported alongside primary results. If counts diverge by more than 30% or verdict tiers systematically differ, the signal is at least partly a translation artifact; this is the hard outcome, and the honest response is to publish the finding, identify which markers are most translation-sensitive, and either restrict claims to translation-robust markers or restructure the lexicon to reduce the artifact contribution.

I want to be candid about my prior expectation. I expect the test, when implemented, to fall in the middle category: substantial cross-translation agreement, with some translation sensitivity concentrated in markers that depend on specific UTV vocabulary choices (the Identity marker is the most likely candidate, since the UTV's rendering of divine names differs from common-translation conventions). If that expectation is correct, the published cross-translation deltas will sharpen rather than weaken the SHIELD claims by demonstrating which signals are robust and which are translation-dependent. If my expectation is wrong in either direction — if agreement is much higher or much lower than I anticipate — the published result will be the more interesting one, and I will report it as I find it.

The translation-invariance test is, in some respects, the most consequential single experiment the SHIELD Protocol can perform on itself. Once published, it will permanently bound the strength of every

future claim. A pattern that survives cross-translation testing is a pattern about scripture; a pattern that does not is a pattern about the UTV. Both are interesting things to know. The first is the more interesting; the second is the more honest.

Chapter 10

What This Release Does Not Settle

It is the responsibility of every protocol release to disclose its limits as plainly as it discloses its capabilities. SHIELD 15.2 does substantially more than 15.1 to defend the empirical core of the methodology against methodological attack. It does nothing at all to address the deeper question, which is not methodological but metaphysical.

The SHIELD Protocol measures whether observed clustering in randomly drawn samples of scriptural text departs from chance. When triple-convergent significance is achieved under hardened conditions — published lexicon, transparent weights, BH-and-BY-and-e-BH agreement, SCS-Pure as headline, externally timestamped pre-registration, translation-invariant signal — the protocol can claim, with substantial confidence, that something non-random is happening. The protocol cannot, by any conceivable hardening, identify what that something is.

The leap from "non-random clustering" to "divine communication" is a theological act. It is not a statistical act. No methodological upgrade collapses the distance between the two, and any reader of this volume who hopes that one will is hoping for a category of evidence the protocol does not produce and cannot produce. What the protocol produces is, at its best, an unimpeachable empirical finding about clustering. The interpretation of that finding remains the interpretive responsibility of the reader, the practitioner, and the community of inquiry within which the finding is received.

This is, in fact, the correct division of labor between the statistical engine and the theologian. The protocol's job is to produce findings that are

robust to every available challenge of method. The theologian's job, having received such findings, is to interpret what they mean in light of broader convictions about the nature of scripture, the nature of providence, the nature of the relationship between the divine and the textual. SHIELD 15.2 brings the protocol substantially closer to discharging its job. It does not, and cannot, discharge the theologian's job. The interpretive layer remains where it has always been: in the reader's soul, before the text, with whatever framework of conviction the reader brings.

This is not a concession that the SHIELD Protocol's findings are unimportant or that they cannot bear theological weight. It is a clarification of the kind of weight they can bear. A finding of robust statistical anomaly does not prove that the source of the anomaly is divine. It does establish that any account of the anomaly that requires it to be a coincidence has become more expensive, because the anomaly is robust to the kinds of method-of-analysis challenges that usually dissolve apparent coincidences. The believer is entitled to take the established robustness as one input among many to a theological position; the skeptic is entitled to seek some non-divine account of the established robustness; the agnostic is entitled to register the finding without yet committing to either side of the interpretive question. The protocol owes each of them the same thing: an empirical finding that can stand to be examined.

There is one further limit worth naming explicitly. The hardening described in this volume addresses the methodological vulnerabilities that an adversarial reviewer can articulate in statistical terms. There are other categories of vulnerability the audit did not address, because they fall outside the statistical frame: questions of whether the practice of cryptographically-seeded scripture-drawing is itself a legitimate spiritual exercise, questions of whether the framing of the operator's inquiry already constrains the answer in ways the protocol cannot detect, questions of whether the entire enterprise mistakes a structural pattern in the canon for a personal address from its Author. These questions are real, they are

unanswered, and SHIELD 15.2 does not pretend to answer them. They are appropriately addressed in other registers — in theology, in spiritual direction, in the long conversation of the church with itself. The protocol's contribution to those conversations is bounded by what the protocol can do, which is statistics. It is not bounded by what statistics can do, which is everything statistics has ever done. The line between the two is the line this chapter is drawing.

Chapter 11

A Note to Skeptics, A Note to Believers

To the skeptical reader, who came to this book expecting to find pseudoscience and may feel surprised to encounter what looks like methodology, a brief direct word.

The SHIELD Protocol is not asking you to share the theological framework within which it was developed. It is asking you to evaluate its empirical claims on the merits of its method. SHIELD 15.2 has done what I could think to do to make those merits assessable: publish the lexicon, publish the weights, publish the formulas, publish the dependency structure, decouple the statistical headline from the interpretive one, anchor the pre-registration externally, slate the translation-invariance test for execution. If you find a flaw I have not anticipated, the appropriate response is to publish the flaw; the protocol's development has, from its earliest versions, depended on exactly that kind of external input, and the response to the present audit is the most substantial example to date of how I treat such input. If you find no flaw and the empirical findings nonetheless persist, you are entitled to seek a non-theological account of the findings, and I would welcome the account. What I would ask, in return for the labor of hardening, is that the empirical findings be engaged on their statistical merits before they are dismissed on theological associations. That is the courtesy any methodology asks of its critics; it is the courtesy SHIELD 15.2 attempts to earn.

To the believing reader, who came to this book expecting confirmation of what is already held and may feel disappointed by what reads like academic concession, a different but equally direct word.

The hardening described in this volume is not a retreat from the protocol's claims. It is a sharpening of them. A claim that has survived BH, BY, and e-BH simultaneously, against a published lexicon, with weights disclosed, with the interpretive layer cleanly separated, externally pre-registered, translation-invariant where translation invariance has been tested, is a stronger claim than a claim that has survived only BH against a private lexicon with fused interpretive weights. The smaller, harder-to-dismiss number that emerges from the hardened protocol is a better witness to whatever the protocol is witnessing than the larger, easily-dismissed number it replaces. If the LFE practice is genuinely receiving the kind of communication some practitioners have understood it to receive, the hardened protocol will register that communication more credibly, not less.

The believer's temptation, when reading methodological self-critique, is to feel that the author has lost confidence in the work. The opposite is true. The author who undertakes methodological self-critique has greater confidence in the work, not less, because the author trusts that the work can survive the critique. A weak claim must be defended; a strong claim can be hardened. SHIELD 15.2 is a hardening, not a defense.

There is a final temptation for both kinds of reader, which is to suppose that the work of methodology can settle the work of faith one way or the other. It cannot. The methodology, however hardened, can produce empirical findings. The findings, however robust, must still be interpreted. The interpretation is a work for which methodology can supply inputs but cannot supply conclusions, and a methodology that pretended otherwise would be a methodology that had overstepped its own competence. SHIELD 15.2 is, more than any prior release, a protocol that knows what it can and cannot do. That knowledge is the foundation on which any future use of the protocol must rest.

Selah.

Appendix A

The Full Decision Hierarchy v2

The Decision Hierarchy is the rule by which SHIELD verdicts are assigned to published findings. SHIELD 15.2 supersedes the 15.1 hierarchy with the following structure.

Tier 1: Robust Significance

All three of BH q-value, BY q-value, and e-BH cross their respective significance thresholds ($q \leq 0.05$ for the two q-values; the e-process exceeds the pre-registered threshold, typically $e \geq 20$). Citation grade: A. Publication recommendation: cite without caveat as triple-convergent significance.

Tier 2: Dependence-Robust Significance

BY and e-BH both cross threshold; BH may or may not. The result is valid under arbitrary dependence among test statistics and under sequential interpretation. Citation grade: B. Publication recommendation: cite as dependence-robust significance with brief note on the dependence-robust qualification.

Tier 3: Anytime-Valid Significance

Only e-BH crosses threshold. The result is valid under arbitrary stopping rules and sequential interpretation; it has not been confirmed under fixed-sample procedures. Citation grade: C. Publication recommendation: cite as anytime-valid significance with explicit note on the sample-size-adaptive

interpretation.

Tier 4: Suggestive Only

Only BH crosses threshold. The result is suggestive but is not robust to the dependence and stopping-rule concerns that BY and e-BH address. Citation grade: D. Publication recommendation: do not cite in publications; if reported, report explicitly as suggestive-only and not as a formal significance claim.

Tier 5: No Significance

None of the three procedures crosses threshold. The result is null. Citation grade: F. Publication recommendation: report as null finding; the absence of significance is itself a publishable result and should be reported as such, in service of avoiding the file-drawer effect.

The Triple-Convergence Standard is the new default reporting convention for all SHIELD publications produced under specification version 15.2 or later. Publications produced under earlier specifications retain their original tier assignments but should, when re-cited, include a footnote noting that the assignment predates the Triple-Convergence Standard and may downgrade under re-analysis.

Appendix B

Implementation Manifest for SHIELD 15.2

The following artifacts are released with this specification and constitute the complete 15.2 hardening package.

Specification Documents

- **SHIELD_15.2_Specification.md** — the full updated specification, incorporating the Triple-Convergence Standard, the SCS-Pure metric, the published CCI weight schedule, and the GDI formula disclosure.
- **SHIELD_15.2_Hardening_Patch.md** — the detailed change record documenting each modification relative to 15.1, with implementation notes and rationale for each change.

Lexicon and Co-Dependency

- **MARKER_LEXICON_v1.0.0.md** — the published, hashed, version-locked marker lexicon defining the I, P, G, T, and SAM tag categories, with primary terms, secondary terms, and negative-context exclusions for each category, and the required tagging-method disclosure.
- **MARKER_CODEPENDENCY_v1.0.0.md** — the published φ correlation matrix for marker pairs across UTV-34211, with raw cell counts permitting independent re-computation under alternative co-dependency measures.

Pre-Registration Infrastructure

- **preregistration_manifest_schema.json** — the JSON schema for the SHIELD pre-registration manifest, supporting validation of any submitted manifest against the specification.
- **sample_preregistration_manifest.json** — a worked example of a complete pre-registration manifest, showing all required fields populated.

Implementation Modules

- **shield_15_2_modules.js** — drop-in JavaScript modules for the LFE Workstation v4 implementing the Benjamini-Yekutieli FDR adjustment, the SCS-Pure index computation, the φ co-dependency matrix, the Triple-Convergence verdict assignment, the pre-registration manifest builder, and the SHA-256 hash convenience.

Release Documentation

- **SHIELD_Protocol_15_2_A_Methodological_Hardening.epub** — this volume.

All artifacts are released simultaneously and constitute a single coherent specification. Implementations of the SHIELD Protocol at specification version 15.2 must conform to all of them; implementations conforming to a subset should declare themselves as "partial 15.2 conformance" with explicit notation of which artifacts are not implemented.

About the Author

Dallas James is the translator of the Urim Thummim Version (UTV), a Masoretic-based English translation of the biblical canon containing 34,211 verses across the Dominion Order Canon. He is the principal researcher at Hearts2Fathers Publishing and the Royal Levitical Priesthood Ministry, the developer of the SHIELD Protocol for statistical analysis of pre-registered scriptural draws, and the author of more than one hundred and nineteen books on biblical translation, prophecy, and the Life Force Experiment. He has approximately thirty years of biblical translation experience, working from the Hebrew Masoretic Text and (separately, for the H2F translation) from Thomson's 1808 Septuagint.

His theological research centers on the Attached Agency Model of pneumatology, the Royal Levitical Priesthood as a structural pattern in scripture, the Dominion Triad convergence (Identity, Inheritance, Priesthood), and the Shroud of Turin as forensic witness to the transphysical resurrection. He operates a YouTube channel under the Royal Levitical Priesthood banner with more than six hundred videos, and additional content on Rumble connected to his ongoing translation and LFE research.

He resides in the Pacific Northwest with his wife and remains an active translator, researcher, and publisher.

About Hearts2Fathers Publishing

Hearts2Fathers Publishing is the publication imprint of the Royal Levitical Priesthood Ministry, founded by Dallas James for the dissemination of original biblical translation, theological research, and methodologically transparent results from the Life Force Experiment.

The press publishes both the Urim Thummim Version (UTV) and the Hearts2Fathers Septuagint (H2F) translations, alongside the ongoing *Transphysical Testimony* series documenting Life Force Experiment draw series on the Shroud of Turin and related subjects, and methodological volumes such as the present work.

All publications are released through standard ebook distribution channels and remain in the catalog of the press in perpetuity.